

WHAT IS CLAIMED IS:

- 1 1. A system for extracting information from articles and for storing the extracted  
2 information in a frame-based knowledge representation, the system comprising:
  - 3 an article selection unit, for selecting and prioritizing articles from which  
4 information will be extracted;
  - 5 an information extraction unit coupled to and in communication with the  
6 article selection unit, which receives selected articles from the article selection  
7 unit and extracts information from the selected article according to pre-defined  
8 information extraction protocols;
  - 9 a knowledge representation management unit, coupled to and in  
10 communication with the information extraction unit for determining if the  
11 extracted information has been both properly extracted and formatted for storage  
12 in the frame-based knowledge representation;
  - 13 an information storage unit coupled to and in communication with the  
14 knowledge representation management unit for storing the information in the  
15 representation if it has been properly extracted and formatted and for responding  
16 to inquiries regarding the stored representation; and
  - 17 a query management and information display unit, coupled to and in  
18 communication with the information storage unit for responding to user inquiries  
19 for information stored in the information storage unit, for retrieving information  
20 from the information storage unit in response to the queries and for displaying the  
21 retrieved information.
- 22
- 23 2. The system of claim 1 wherein the information extraction unit and knowledge  
24 representation management unit are combined.
- 25
- 26 3. The system of claim 1 wherein at least the information extraction unit and the  
27 knowledge representation management unit are geographically widely separated, with the  
28 respective units being located wherever the functions of the respective units can be  
29 performed at the lowest cost.
- 30
- 31 4. A method for constructing a frame-based knowledge representation, the method  
32 comprising the steps of:

33           selecting articles to serve as an information source for the knowledge  
34 representation;  
35           extracting information from the selected articles and formatting it for storage in  
36 the knowledge representation;  
37           verifying that the information extracted from the selected articles is correct and  
38 that it has been placed in the correct format; and  
39           storing the formatted information in the knowledge representation.

40

41   5. The method of claim 4 wherein the extracting information step is performed by  
42 knowledge extraction personnel and the verifying step is performed by quality control  
43 personnel.

44

45   6. The method of claim 5 wherein both the extracting step and verifying step are  
46 performed by the same person, which person has been qualified by a predetermined  
47 procedure to perform both steps simultaneously.

48

49   7. The method of claim 4 wherein at least the steps of extracting and verifying occur in  
50 geographically separated locations.

51

52   8. The method of claim 7 wherein the geographically separate locations are chosen based  
53 upon the cost of performing the respective steps of extracting and verifying, the lowest  
54 cost location for each step being selected.

55